
Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSEARCH protein - protein database search, using Smith-Waterman algorithm
Run on: Sat May 13 09:35:48 2000; MasPar time 3.96 Seconds
Tabular output not generated.

Title: >US-09-331-631-1
Description: (74-116) from US09331631.pep (3 of 5)
Perfect Score: 344
Sequence: 1 NOEDPOTCCOCCRRCOESGPRQOQYCORCKEICEEEY 43
Scoring table: PAM 150
Gap 11
Searched: 188963 seqs, 23666106 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: a-geneseq35
1:geneseqp
Statistics: Mean 22.277; Variance 95.380; scale 0.234

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	344	100.0	666	1	Macadamia integrifolia	7.91e-24
2	342	99.4	666	1	Macadamia integrifolia	1.23e-23
3	337	98.0	625	1	Macadamia integrifolia	3.73e-23
4	162	47.1	525	1	Macadamia integrifolia	1.03e-06
5	162	47.1	566	1	Macadamia integrifolia	1.03e-06
6	129	37.5	590	1	Macadamia integrifolia	1.03e-06
7	113	32.8	28	1	Macadamia integrifolia	2.22e-02
8	111	32.3	593	1	Macadamia integrifolia	3.29e-02
9	105	30.5	33	1	Macadamia integrifolia	1.07e-01
10	96	27.9	35	1	Macadamia integrifolia	6.11e-01
11	85	24.7	637	1	Macadamia integrifolia	4.91e+00
12	83	24.1	106	1	Macadamia integrifolia	7.13e+00
13	83	24.1	106	1	Macadamia integrifolia	7.13e+00
14	82	23.8	626	1	Macadamia integrifolia	8.58e+00
15	80	23.3	441	1	Macadamia integrifolia	1.24e+01
16	80	23.3	441	1	Macadamia integrifolia	1.24e+01
17	80	23.3	919	1	Macadamia integrifolia	1.24e+01
18	80	23.3	919	1	Macadamia integrifolia	1.24e+01
19	77	22.4	614	1	Macadamia integrifolia	2.15e+01
20	77	22.4	614	1	Macadamia integrifolia	2.15e+01
21	77	22.4	918	1	Macadamia integrifolia	2.15e+01
22	76	22.1	439	1	Macadamia integrifolia	2.58e+01
23	76	22.1	440	1	Macadamia integrifolia	2.58e+01

ID	Score	Query Match	Length	ID	Description	Pred. No.
24	76	22.1	440	1	Macadamia integrifolia	2.58e+01
25	76	22.1	910	1	Macadamia integrifolia	2.58e+01
26	76	22.1	1058	1	Macadamia integrifolia	2.58e+01
27	76	22.1	1058	1	Macadamia integrifolia	2.58e+01
28	76	22.1	1308	1	Macadamia integrifolia	2.58e+01
29	76	22.1	1308	1	Macadamia integrifolia	2.58e+01
30	74	21.5	805	1	Macadamia integrifolia	3.71e+01
31	73	21.2	768	1	Macadamia integrifolia	4.44e+01
32	72	20.9	468	1	Macadamia integrifolia	5.31e+01
33	71	20.6	405	1	Macadamia integrifolia	6.35e+01
34	71	20.6	509	1	Macadamia integrifolia	6.35e+01
35	71	20.6	529	1	Macadamia integrifolia	6.35e+01
36	70	20.3	518	1	Macadamia integrifolia	7.39e+01
37	69	20.1	204	1	Macadamia integrifolia	9.06e+01
38	69	20.1	402	1	Macadamia integrifolia	9.06e+01
39	69	20.1	898	1	Macadamia integrifolia	9.06e+01
40	68	19.8	145	1	Macadamia integrifolia	1.08e+02
41	68	19.8	145	1	Macadamia integrifolia	1.08e+02
42	68	19.8	190	1	Macadamia integrifolia	1.08e+02
43	68	19.8	220	1	Macadamia integrifolia	1.08e+02
44	68	19.8	1003	1	Macadamia integrifolia	1.08e+02
45	68	19.8	1786	1	Macadamia integrifolia	1.08e+02

ALIGNMENTS

ID	Score	Query Match	Length	ID	Description	Pred. No.
1	344	100.0	666	1	Macadamia integrifolia	7.91e-24
2	342	99.4	666	1	Macadamia integrifolia	1.23e-23
3	337	98.0	625	1	Macadamia integrifolia	3.73e-23
4	162	47.1	525	1	Macadamia integrifolia	1.03e-06
5	162	47.1	566	1	Macadamia integrifolia	1.03e-06
6	129	37.5	590	1	Macadamia integrifolia	1.03e-06
7	113	32.8	28	1	Macadamia integrifolia	2.22e-02
8	111	32.3	593	1	Macadamia integrifolia	3.29e-02
9	105	30.5	33	1	Macadamia integrifolia	1.07e-01
10	96	27.9	35	1	Macadamia integrifolia	6.11e-01
11	85	24.7	637	1	Macadamia integrifolia	4.91e+00
12	83	24.1	106	1	Macadamia integrifolia	7.13e+00
13	83	24.1	106	1	Macadamia integrifolia	7.13e+00
14	82	23.8	626	1	Macadamia integrifolia	8.58e+00
15	80	23.3	441	1	Macadamia integrifolia	1.24e+01
16	80	23.3	441	1	Macadamia integrifolia	1.24e+01
17	80	23.3	919	1	Macadamia integrifolia	1.24e+01
18	80	23.3	919	1	Macadamia integrifolia	1.24e+01
19	77	22.4	614	1	Macadamia integrifolia	2.15e+01
20	77	22.4	614	1	Macadamia integrifolia	2.15e+01
21	77	22.4	918	1	Macadamia integrifolia	2.15e+01
22	76	22.1	439	1	Macadamia integrifolia	2.58e+01
23	76	22.1	440	1	Macadamia integrifolia	2.58e+01

FT Protein 29. .666
 /note="mature protein"
 PN WO9827805-A1.
 PD 02-JUL-1998.
 FE 22-DEC-1997: AU0874.
 PR 20-DEC-1996: AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NI, Goulter KC, Green JL, Mannes JM, Marcus JP;
 DR WPI: 98-377279/32.
 N-PSDB: V42311.
 PI Novel anti-microbial protein from e.g. Macadamia integrifolia -
 useful for controlling microbial infestations of plants or mammals
 PS Claim 1: Page 39-41; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can
 be used to control microbial infestations in plants and mammalian
 CC animals.
 SQ Sequence 666 AA;

Query Match 99.4%; Score 342; DB 1; Length 666;
 Best Local Similarity 95.3%; Pred. No. 1.23e-23;
 Matches 41; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 74 NOEDPOTECOCRCRCROESGPRQOYQCRCKEICEEEY 116
 QY 74 NOEDPOTECOCRCRCROESGPRQOYQCRCKEICEEEY 116

RESULT 3
 ID W62830 standard; Protein; 625 AA.
 AC W62830:
 DT 27-OCT-1998 (first entry)
 DE Macadamia integrifolia antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Macadamia integrifolia.
 FH Key Location/Qualifiers
 FT Peptide 1. .28
 FT /note="signal peptide"
 FT 29. .666
 FT Protein /note="mature protein"
 PN WO9827805-A1.
 PD 02-JUL-1998.
 FE 22-DEC-1997: AU0874.
 PR 20-DEC-1996: AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NI, Goulter KC, Green JL, Mannes JM, Marcus JP;
 DR WPI: 98-377279/32.
 N-PSDB: V42311.
 PI Novel anti-microbial protein from e.g. Macadamia integrifolia -
 useful for controlling microbial infestations of plants or mammals
 PS Claim 1: Page 43-45; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can
 be used to control microbial infestations in plants and mammalian
 CC animals.
 SQ Sequence 625 AA;

Query Match 98.0%; Score 337; DB 1; Length 625;
 Best Local Similarity 97.7%; Pred. No. 3.73e-23;
 Matches 42; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 33 NOEDPOTECOCRCRCROESGPRQOYQCRCKEICEEEY 75
 QY 74 NOEDPOTECOCRCRCROESGPRQOYQCRCKEICEEEY 116

RESULT 4
 ID W62831 standard; Protein; 525 AA.
 AC W62831:
 DT 27-OCT-1998 (first entry)
 DE Theobroma cacao antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Theobroma cacao.
 PN WO9827805-A1.
 PD 02-JUL-1998.

PF 22-DEC-1997: AU0874.
 PR 20-DEC-1996: AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NI, Goulter KC, Green JL, Mannes JM, Marcus JP;
 DR WPI: 98-377279/32.
 PI Novel anti-microbial protein from e.g. Macadamia integrifolia -
 useful for controlling microbial infestations of plants or mammals
 PS Claim 1: Page 47-49; 96pp; English.
 CC The sequence is that of an antimicrobial protein which can
 be used to control microbial infestations in plants and mammalian
 CC animals.
 SQ Sequence 525 AA;

Query Match 47.1%; Score 162; DB 1; Length 525;
 Best Local Similarity 50.0%; Pred. No. 1.03e-06;
 Matches 20; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

DB 78 EEELORQYQCCGRCRCROESGPRQOYQCRCKEICEEEY 117
 QY 75 QEDPOTECOCRCRCROESGPRQOYQCRCKEICEEEY 114

RESULT 5
 ID R20181 standard; Protein; 566 AA.
 AC R20181:
 DT 16-APR-1992 (first entry)
 DE Sequence encoded by 67 kD T. cacao protein cDNA.
 KW Cocoa; flavour; vicillin; seed storage protein.
 OS Theobroma cacao.
 PN WO9119801-A.
 PD 26-DEC-1991.
 PF 07-JUN-1991; G00914.
 PR 11-JUN-1990; GB-013016.
 PA (MRC) MARS UK LTD.
 PI Spencer ME, Hodge R, Deakin EA, Ashton S;
 DR WPI: 92-024418/03.
 N-PSDB: Q20377.
 PI Recombinant cocoa proteins - are responsible for flavour in cocoa
 PT beans and produced in large quantities using yeast and bacterial
 PT expression vectors
 PS Claim 4: Fig 2; 59pp; English.
 CC The inventors claim a 67 kD and 31 kD T. cacao protein, and
 CC fragments, and encoding DNAs. The 47 kD and 31 kD proteins are
 CC derived from the 67 kD precursor. T. cacao protein cDNA was
 CC detected in a cDNA library prepared from immature cocoa beans RNA
 CC using a probe based on the AA sequence of a CNBR peptide common to
 CC the 47 kD and 31 kD polypeptides. Homology searches revealed close
 CC homologies between the 67 kD polypeptide and the vicillins, which are
 CC seed storage proteins.
 SQ Sequence 566 AA;

Query Match 47.1%; Score 162; DB 1; Length 566;
 Best Local Similarity 50.0%; Pred. No. 1.03e-06;
 Matches 20; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

DB 78 EEELORQYQCCGRCRCROESGPRQOYQCRCKEICEEEY 117
 QY 75 QEDPOTECOCRCRCROESGPRQOYQCRCKEICEEEY 114

RESULT 6
 ID W62832 standard; Protein; 590 AA.
 AC W62832:
 DT 27-OCT-1998 (first entry)
 DE Gossypium hirsutum antimicrobial protein.
 KW antimicrobial protein; infestation; control.
 OS Gossypium hirsutum.
 PN WO9827805-A1.
 PD 02-JUL-1998.
 PF 22-DEC-1997: AU0874.
 PR 20-DEC-1996: AU-004275.
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
 PI Bower NI, Goulter KC, Green JL, Mannes JM, Marcus JP;


```
RESULT 11
ID W62837 standard; Protein: 637 AA.
AC W62837;
DE 27-OCT-1998 (first entry)
KW Hordium vulgare antimicrobial protein.
OS Hordium vulgare.
PN W09827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997: AU0874.
PR 20-DEC-1996: AU-004275.
PA (RER-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NJ, Goulter KC, Green JL, Mannes JM, Marcus JP;
DR WPI: 98-37279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PS useful for controlling microbial infestations of plants or mammals
CC Claim 1: Page 60-62; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 637 AA;

Query Match 24.7%; Score 85; DB 1; Length 637;
Best Local Similarity 42.4%; Pred. NO. 4.91e+00;
Matches 14; Conservative 7; Mismatches 7; Indels 5; Gaps 3;

DB 42 OOCVORCRQ-ER-PR---YSHARCVOECRDDQ 69
QY 83 QQCQRCKRQDESGFRQOQCQRCKCECEEE 115

RESULT 12
ID R91706 standard; Protein: 106 AA.
AC R91706;
DE 17-NOV-1996 (first entry)
KW Acetabularia; HPOMAP; NamMAP; AcenMAP; AdunMAP; anticoagulant;
KW nematode-extracted anticoagulant protein; serine protease;
KW nematode; thrombosis; parasitic worm.
OS Ancylostoma caninum.
PN W09612021-A2.
PD 25-APR-1996.
PF 17-OCT-1995; U13231.
PR 18-OCT-1994; US-326110.
PR 05-JUN-1995; US-486399.
PR 05-JUN-1995; US-461965.
PR 05-JUN-1995; US-486397.
PR 05-JUN-1995; US-465380.
PA (CORV-) CORVAS INT INC.
PI Bergum PW, Ganssemans YGJ, Jespers LS, Laroche YR;
PI Lauwereys MJ, Messens JH, Moyle M, Stanssens PEH;
PI Vlaek GP;
DR WPI: 96-222007/22.
DR N-PSDB: T12951.
PT Proteins with anticoagulant and/or serine protease inhibitory
PT activity - isolated from nematodes and useful to inhibit blood
PT coagulation
PS Claim 221; Fig 13B; 243pp; English.
CC Proteins with anticoagulant and/or serine protease inhibitory
CC activity, isolated from nematodes, are useful to inhibit blood
CC coagulation. The proteins can be added to blood collection tubes
CC defining the collection of mammalian plasma. They are also useful
CC to prevent or inhibit thrombosis, and may be given alone or in
CC combination with other therapeutic or in vivo diagnostic agents.
CC The proteins can serve as immunogens to raise antibodies for use in
CC the diagnosis and identification of NBP concn. levels in biological
CC fluids, e.g. to detect mammalian infection with a parasitic worm.
CC They can also be used as immunogens in prophylactic and therapeutic
CC vaccines against parasitic worm infection. The proteins may
CC double the clotting time of human plasma in prothrombin time assays
CC when present at 10-50 nMol, and double the clotting time of human
CC plasma in activated partial thrombin time assays when present
```

```
CC at 10-100 nMol.
CC The anticoagulant proteins are pref. derived from
CC Ancylostoma caninum, A. ceylanicum, A. duodenale, Necator
CC americanus or Heligmosomoides polygyrus.
CC The proteins pref. have 2 NBP domains and specifically inhibit
CC the catalytic activity of the factor VIIa/TF complex in the
CC presence of factor Xa or a catalytically inactive factor Xa deriv.,
CC do not specifically inhibit the activation of factor VIIa in the
CC absence of TF and do not specifically inhibit prothrombinase.
SQ Sequence 106 AA;

Query Match 24.1%; Score 83; DB 1; Length 106;
Best Local Similarity 56.3%; Pred. NO. 7.13e+00;
Matches 9; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

DB 39 CERCKIETSEEDDY 54
QY 102 CORCK-ECICEEEY 116

RESULT 13
ID R91705 standard; Protein: 107 AA.
AC R91705;
DE 17-NOV-1996 (first entry)
KW Acetabularia; HPOMAP; NamMAP; AcenMAP; AdunMAP; anticoagulant;
KW nematode-extracted anticoagulant protein; serine protease;
KW nematode; thrombosis; parasitic worm.
OS Ancylostoma caninum.
PN W09612021-A2.
PD 25-APR-1996.
PF 17-OCT-1995; U13231.
PR 18-OCT-1994; US-326110.
PR 05-JUN-1995; US-486399.
PR 05-JUN-1995; US-461965.
PR 05-JUN-1995; US-486397.
PR 05-JUN-1995; US-465380.
PA (CORV-) CORVAS INT INC.
PI Bergum PW, Ganssemans YGJ, Jespers LS, Laroche YR;
PI Lauwereys MJ, Messens JH, Moyle M, Stanssens PEH;
PI Vlaek GP;
DR WPI: 96-222007/22.
DR N-PSDB: T12951.
PT Proteins with anticoagulant and/or serine protease inhibitory
PT activity - isolated from nematodes and useful to inhibit blood
PT coagulation
PS Claim 221; Fig 13A; 243pp; English.
CC Proteins with anticoagulant and/or serine protease inhibitory
CC activity, isolated from nematodes, are useful to inhibit blood
CC coagulation. The proteins can be added to blood collection tubes
CC defining the collection of mammalian plasma. They are also useful
CC to prevent or inhibit thrombosis, and may be given alone or in
CC combination with other therapeutic or in vivo diagnostic agents.
CC The proteins can serve as immunogens to raise antibodies for use in
CC the diagnosis and identification of NBP concn. levels in biological
CC fluids, e.g. to detect mammalian infection with a parasitic worm.
CC They can also be used as immunogens in prophylactic and therapeutic
CC vaccines against parasitic worm infection. The proteins may
CC double the clotting time of human plasma in prothrombin time assays
CC when present at 10-50 nMol, and double the clotting time of human
CC plasma in activated partial thrombin time assays when present
CC at 10-100 nMol.
CC The anticoagulant proteins are pref. derived from
CC Ancylostoma caninum, A. ceylanicum, A. duodenale, Necator
CC americanus or Heligmosomoides polygyrus.
CC The proteins pref. have 2 NBP domains and specifically inhibit
CC the catalytic activity of the factor VIIa/TF complex in the
CC presence of factor Xa or a catalytically inactive factor Xa deriv.,
CC do not specifically inhibit the activation of factor VIIa in the
CC absence of TF and do not specifically inhibit prothrombinase.
SQ Sequence 107 AA;

Query Match 24.1%; Score 83; DB 1; Length 107;
```

Search completed: Sat May 13 09:35:56 2000
Job time : 8 secs.

